

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G01N 33/53, 33/536, 23/00, 33/00, C12Q 1/37, C07C 229/00, C07D 311/82	A1	(11) International Publication Number: WO 00/16100 (43) International Publication Date: 23 March 2000 (23.03.00)
(21) International Application Number: PCT/US99/21177 (22) International Filing Date: 13 September 1999 (13.09.99) (30) Priority Data: 60/099,847 11 September 1998 (11.09.98) US (71) Applicant: THERMAPHORE SCIENCES, INC. [US/US]; 12 East 49th Street, New York, NY 10017 (US). (72) Inventor: WOODS, Virgil, L., Jr.; 2920 Union Street, #103, San Diego, CA 92103 (US). (74) Agents: GISOLFI, Ann, L. et al.; Pennie & Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036 (US).		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i>
(54) Title: METHODS FOR IDENTIFYING HOT-SPOT RESIDUES OF BINDING PROTEINS AND SMALL COMPOUNDS THAT BIND TO THE SAME (57) Abstract The present invention provides methods of identifying hot-spot residues for one or both members of a receptor-ligand complex of interest. Further provided are methods of using receptor hot-spot residues to identify compounds that functionally bind a receptor in a manner that mimics the binding of a known ligand for the receptor.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Kazakhstan	PT	Portugal		
CU	Cuba	LC	Saint Lucia	RO	Romania		
CZ	Czech Republic	LI	Liechtenstein	RU	Russian Federation		
DE	Germany	LR	Liberia	SD	Sudan		
DK	Denmark			SE	Sweden		
EE	Estonia			SG	Singapore		

**METHODS FOR IDENTIFYING HOT-SPOT RESIDUES OF BINDING PROTEINS
AND SMALL COMPOUNDS THAT BIND TO THE SAME**

5 1. **FIELD OF THE INVENTION**

The present invention relates to methods of identifying hot-spot residues of a member of a receptor-ligand binding pair of interest. The invention further provides methods of using information about receptor hot-spots for a receptor-ligand pair of interest to guide the identification of compounds that functionally bind to such regions of the
10 receptor in a manner that mimics the ligand.

2. **BACKGROUND OF THE INVENTION**

Protein-protein interactions such as those observed in many receptor-ligand complexes are typically mediated by large binding surfaces comprising ten to thirty contact
15 amino acid residues on each protein of the complex (Clackson & Wells, 1995, Science 267:383-386). Recently, it has been suggested for some binding pairs that within each binding surface, the predominant contribution to the overall free energy of binding of the complex is due to only a few residues within each binding surface. (Clackson & Wells, 1995, *supra*).

20 For example, recent studies by Wells and collaborators have characterized the binding surfaces of the complex formed between human growth hormone and its cellular receptor. (Clackson & Wells, 1995, *supra*; Pearce et al., 1996, Biochemistry 35:10300-10307; Wells, 1996, Proc. Natl. Acad. Sci. USA 93:1-6). X-ray crystallographic studies had shown that approximately thirty amino acid side chains of the hormone contact
25 approximately thirty amino acid side chains of the receptor. An array of mutant human growth hormone proteins and mutant receptor proteins were prepared to estimate the energetic contribution of each amino acid of the contact surface to the overall binding interaction. The study measured the binding affinities of various complexes between mutant hormones and native receptors and also between native hormones and mutant
30 receptors.

Surprisingly, only a small number of amino acid residues within the extensive binding surfaces accounted for most of the binding energy. Overall, fewer than half of the contact residues contributed measurably to the binding interaction. Eight receptor residues and eleven hormone residues accounted for over 75% of the free energy of binding of the
35 hormone-receptor complex.

Within each group of residues on either binding pair, a limited number of "hot-spot" amino acids contributed most to the total free energy of binding. Surrounding each "hot-spot" was a group of several amino acids that contributed at lesser levels, and this group was in turn followed by a larger set of residues that contributed at even lower levels to the total free energy of binding. Most significantly, the "hot-spot" residues on the receptor binding surface directly contacted the complementary "hot-spot" amino acids on the hormone's binding surface. Subsequent studies in other systems have also shown that the total binding affinity between a receptor and a ligand is mediated mainly by small and complementary sets of "hot-spot" residues within the binding surfaces of the receptor and ligand (Wells, 1996, Science 273:449-450; Smith-Gill, 1994, Res. Immunol. 145:67-70, 1994). Some systems even suggest that multiple, non-contiguous "hot-spots" are possible.

While the methods of Wells and others can provide detailed information about residues involved in binding interactions for a variety of receptor-ligand pairs, they suffer from serious drawbacks. For example, for a given receptor-binding protein pair, large panels of mutant proteins must be prepared in order to identify the thermodynamically significant or "hot-spot" residues. To do so, each residue of the receptor or ligand must be mutated individually to measure its contribution to the overall free energy of binding.

Ideally, when a three-dimensional structure of the receptor-ligand complex is available, the mutations can be limited to only those amino acids comprising the binding surfaces of each member of the pair, approximately thirty residues for each number in a typical interaction. Thus, even in such an "ideal" situation, where information about both the receptor and ligand is desired, a total of about sixty mutant proteins must be prepared (one for each residue in the binding surface of the receptor and one for each residue in the binding surface of the ligand), and the binding energy of each mutant protein for its native binding partner determined.

In a more typical situation where no three-dimensional structure is available for either binding partner, mutants must in principle be prepared for every residue of both proteins. Thus, even in an ideal situation the methods are labor-intensive and expensive. In a less than ideal situation, the time and expense of preparing the mutants and measuring binding affinities oftentimes would be comparable to that for obtaining a three-dimensional structure of the complex.

In addition, the methods do not necessarily accurately identify the hot-spot residues. Site-directed mutagenesis of individual amino acids does not simply remove an amino acid side chain from a protein. Often, mutation of a residue to, for example, an alanine, introduces local and perhaps global structural perturbations in the mutant protein. These

structural changes may affect the binding interactions between the mutant protein and its ligand. As a consequence, a comparison of the binding affinities of the mutant and native proteins is unlikely to be an accurate measurement of the relative contribution that residue makes to the overall free-energy of binding. In fact, in the studies of Clackson & Wells, the
5 sum of the apparent free energy contributions of the mutated amino acids exceeded the known free energy of binding of the native receptor-native ligand complex by a factor of two (*see*, Clackson & Wells, 1995, *supra*). Clackson & Wells concluded that the excess measured binding energy was due to mutant-induced structural changes.

Thus, it would be highly desirable to have available methods for identifying hot-spot
10 amino acid residues for receptor-ligand pairs which do not suffer from the above-described limitations. In particular, it would be highly desirable to have available methods for identifying receptor-ligand hot-spot residues which are fast and inexpensive, and which further permit the identification of hot-spot residues for a native receptor-ligand pair without having to mutate either the receptor or ligand.

15 Small compounds that bind to, and thereby antagonize (antagonists) or activate (agonists) receptors of therapeutic importance are of considerable commercial value. However, to date the ability to rapidly and easily identify such small compounds, particularly those that are able to interact with protein receptors that have large polypeptide or protein ligands, is limited. In most instances, such compounds are obtained through the
20 laborious process of rational drug design, which usually requires detailed knowledge about the three-dimensional structure of the receptor-ligand co-complex or other structure-function information. Rational drug design generally involves designing compounds based on available structure-function information on a compound-by-compound basis and screening the individual compounds in biological assays to identify those compounds which
25 produce a desired biological activity. Usually, detailed structure-function or three-dimensional structural information for the receptor-compound complex is obtained so that the design hypotheses can be verified.

Typically, the detailed structure-function information necessary to design and assess the compounds is obtained from NMR or x-ray crystallographic studies with the receptor-
30 ligand co-complex. Both of these technologies require large quantities of pure, co-complex, expensive, specialized equipment and highly skilled technicians, making such structural information extremely costly and time consuming to obtain. In addition, while the detailed structural information obtained from these methods can oftentimes identify those residues of each member of the binding pair that are involved in the binding interaction, neither of
35 these methods have been used to identify the individual residues which make significant

contributions to the overall free energy of binding of the receptor-ligand complex; in fact, x-ray crystallography simply cannot generate this detailed information. Moreover, to date, rational drug design methodologies have proven inadequate for designing receptor agonists, which are an increasingly important class of pharmaceutical targets. Induction of receptor
5 activation by agonists requires more than simple receptor binding. Agonist activity usually results from receptor conformational change or receptor subunit oligomerization that is induced by ligand binding. Agonist ligands must be capable of binding to precisely the correct subregion(s) of the receptor and in the correct manner required for induction of these receptor changes. When agonists appropriately bind to receptors, a portion of the binding
10 energy is employed to produce these receptor structural changes. Presently available structure determination technologies (x-ray crystallography, NMR) have great difficulty in identifying and localizing the parts of the receptor that participate in such ligand-induced structural changes, and in identifying the mechanisms by which they are induced by agonist binding.

15 Recently, "non-rational" combinatorial library methodologies have been developed, in part to ameliorate the drawbacks of rational drug design. In these non-rational methods, large libraries of compounds are synthesized and screened for the ability to bind a target receptor of interest. The structures of the identified compounds are elucidated and used as lead compounds for subsequent rounds of library synthesis and screening or as lead
20 structures for rational drug design.

While these methods have the allure of being able to randomly or semi-rationally sample large sectors of "conformation space" in a relatively short period of time, they suffer from many drawbacks. For example, in order to identify a lead compound from a purely random library, millions, and oftentimes trillions, of compounds must be synthesized and
25 screened. Thus, while theoretically pharmaceutical leads can be identified from a purely random library, in practice the library methods are merged with rational design methods to create "focused libraries" that have a higher probability of containing a suitable lead structure. Thus, even these library methodologies rely, to some extent, on structure-function information.

30 Moreover, while compounds which bind a target receptor can be identified, no information is provided about where on the receptor the compound binds. Whether the compound binds the receptor at or near the same site as the receptor's natural ligand is simply unknown. Combinatorial library methodologies are also inadequate for identifying receptor agonists for substantially the same reasons discussed above.

35

The ability to identify where a candidate compound of interest binds a target receptor of interest without having to obtain expensive and laborious high resolution three-dimensional structures of the receptor, ligand or receptor-ligand co-complex would be highly desirable. Of particular interest would be having the ability to identify candidate compounds which bind a receptor at the same site as the receptor's natural ligand or to identify compounds that act as receptor agonists. Current methods for achieving these goals are inadequate. Accordingly, these are objects of the present invention.

3. SUMMARY OF THE INVENTION

These and other shortcomings in the art are overcome by the present invention, which in one aspect provides a method for identifying hot-spot residues for one or both members of a receptor-ligand pair of interest. In the method of the invention, the rates of exchange between individual hydrogens of one or both members of the receptor-ligand pair and solvent hydrogens are determined for the member in both the bound and unbound states. The individual hydrogens are correlated to specific amino acid residues within the primary sequence of the member, thereby providing hydrogen exchange rates for hydrogens on specifically identified amino acid residues comprising the member. From these exchange rates, those residues which individually contribute at least about 10-20% of the overall free energy of binding of the receptor-ligand complex are identified as constituting the "hot-spot" residues for that member of the receptor-ligand pair.

The method can be used with or without the aid of three-dimensional structure or other structure-function information about the particular receptor, ligand or receptor-ligand pair. The only requirement is that the primary amino acid sequence of the binding member of interest be known.

Furthermore, in general, the method does not preclude application to finding hot-spot residues in a non-native, or mutated receptor. Such an approach could find utility in, for example, establishing structure-function relationships, particularly because mutants often have unpredictable conformational changes with respect to the native receptor.

The methods described herein provide significant advantages over currently available methodologies for identifying receptor and/or ligand hot-spot residues. For example, according to the methods of the invention, hot-spot residues can be simply, rapidly and inexpensively identified for each member of a native receptor-ligand complex. Thus, unlike the methods described in the art, which require mutation of the receptor and/or ligand, the method of the invention permits, for the first time, reliable identification of hot-

35

spot residues of receptors and/or ligands in their truly native, biologically relevant conformations.

In another aspect, the invention provides a method for identifying compounds that interact with a target receptor of interest in a manner similar to a known ligand for that receptor; *i.e.*, the invention provides a method of identifying compounds which act as binding mimics of known receptor ligands. According to the method of the invention, receptor hot-spot residues for a receptor-ligand pair of interest are first identified. Candidate compounds are then screened with the receptor to identify those candidate compounds that interact with at least one, and preferably at least a majority or even more, of the receptor hot-spot residues. Those candidate compounds that interact with at least one receptor hot-spot residue are selected as small molecule binding "mimics" of the known ligand.

For any given receptor-ligand pair, the receptor hot-spot residues can be determined by any of several techniques, including, for example, the methods described herein, hydrogen exchange with heavy hydrogen (deuterium or tritium) in conjunction with nuclear magnetic resonance (NMR) or mass spectroscopy, tritium exchange in conjunction with radioactivity measurements and in conjunction with site-directed mutagenesis. The precise binding interactions between the receptor and candidate compounds can also be characterized by any of these methods to identify compounds that "mimic" the essential receptor binding activity of the known ligand.

Each candidate compound may be first assayed for a particular biological activity of interest, such as the ability to bind, antagonize or agonize the receptor, to identify those candidate compounds which exhibit the desired biological activity, and then the active candidate compounds further screened to identify those that interact with a receptor hot-spot residue. Alternatively, the candidate compounds may be directly screened to identify those that interact with a receptor hot-spot residue without first assaying for biological activity.

Selection of the initial candidate compounds to be screened can be guided by a structural model of the ligand or the receptor-ligand complex, by available structure-function information for the receptor-ligand complex, or by completely random processes. The initial candidate compounds can be synthesized individually or can be synthesized using any of the well-known combinatorial library methodologies. Candidate compounds, and in particular combinatorial libraries of candidate compounds, that are first screened for biological activity, particularly binding activity, can be screened in pools in a batch-wise fashion using any of the numerous screening methods described in the art to identify biologically active compounds for further analysis according to the invention.

Ideally, the method of the invention is used iteratively to identify ligand mimics. In the iterative method, the results of one round of screening guide the selection of candidate compounds for the next round of screening. For example, the structures of candidate compounds identified in a first round of hot-spot screening can be determined and these
5 structures used to guide the synthesis of new or modified candidate compounds. After several rounds of synthesis and screening, compounds are identified that precisely mimic the known ligand; *i.e.*, compounds are identified that precisely interact with those receptor hot-spot residues contacted by the known ligand.

The methods of the invention for identifying compounds which bind a receptor
10 provide significant advantages over currently available techniques. Because the preferred methods utilize sensitive hydrogens on the amino acid side-chain or backbone as reporters of local changes in environment, the methods are able to identify compounds that interact with a target receptor of interest at the most biologically significant residues — those receptor residues that contribute most to binding the receptor's natural ligand. Quite
15 significantly, the methods do not require detailed structure-function or three-dimensional structural information for the receptor, ligand, receptor-ligand co-complex or receptor-compound co-complex, though interpretation of results is enhanced by the availability of such information. The only requirement is that the primary amino acid sequence of the receptor of interest be known. Thus, the methods of the invention permit the identification
20 of "ligand mimics", particularly small compound ligand mimics of the essential binding function of the natural ligand, with unprecedented ease.

Moreover, the methods provide insights into the nature of the ligand-binding induced receptor structural changes required for agonist activity. This, combined with the ability to focus small molecule design to known ligand-defined hot-spots allows ready
25 identification of compounds that function as receptor agonists — a feat that is simply unprecedented in the art.

Thus, in a final aspect the invention provides methods of identifying compounds, particularly small organic compounds, which act as agonists for a target receptor of interest. According to the method of the invention, hydrogen exchange techniques as previously
30 described are used to identify three distinct classes of receptor amino acid residues for a receptor-agonist pair of interest:

(i) those that reside in the receptor-agonist binding surface;

(ii) those that participate in conformational changes as a consequence of receptor-agonist binding; and

35

(iii) that participate in binding-induced receptor sub-unit oligomerization. The receptor is then screened with candidate compounds using the hydrogen exchange techniques described herein to identify those which interact with at least one receptor amino acid residue from class (ii) or (iii), above. The candidate compound is considered to
5 interact with one of these receptor residues if the receptor residue has a protection factor of at least one upon formation of a receptor-candidate compound complex. The candidate compounds may be screened for biological activity prior to screening in the methods of the invention, as previously described.

The methods may be used iteratively in a manner similar to that described above to
10 identify compounds which interact with a majority of the receptor amino acid residues of class (ii) or (iii).

3. DEFINITIONS

As used herein, the following terms shall have the following meanings:
15

"Ligand:" refers to any molecule which is capable of binding a receptor. The ligand may be a small organic compound, or a large molecule such as a polypeptide or protein.

"Receptor:" refers to any molecule capable of binding a ligand. Thus, as
20 used herein receptor refers not only to molecules generally recognized as belonging to the class of binding molecules designated as "receptors," such as, for example, receptors for cytokines, growth factors, chemokines, hormone receptors, adhesion receptors, apoptosis receptors, etc., but is also intended to include any molecule which can bind another molecule, including for example, an antibody or an enzyme.

25 The receptor may be identical in primary amino acid sequence to a naturally occurring receptor, or it may be a functionally active fragment, mutant or derivative thereof. The fragment, mutant or derivative may have the same or different binding or other biological characteristics relative to the parental protein.

Receptors of particular interest include integral membrane proteins, as they
30 are difficult to crystallize for study by X-ray diffraction. Receptors too large to study by NMR methods, e.g., those larger than about 50 kDa, are also of special interest, particularly if they cannot be characterized as a composite of two or more separately analyzable domains. Examples of receptors of particular interest include integrins (which are large integral membrane proteins), cell surface receptors for growth factors (including cytokine

35

receptors), "seven-spanners," selectin, and cell surface receptors of the immunoglobulin superfamily (e.g., ICAM-1).

As will be appreciated by those of skill in the art, the designation of "receptor" and "ligand" is somewhat arbitrary. Typically the term "receptor" is conferred upon the larger of the two binding partners, or upon the member of the binding pair which is a protein. For purposes of the present invention, the expressions "receptor" and "ligand" do not carry such limitations; they merely each refer to one member of a pair of binding molecules.

"Protein:" as used herein includes, *mutatis mutandis*, polypeptides, oligopeptides and derivatives thereof, including, by way of example and not limitation, glycoproteins, lipoproteins, phosphoproteins and metaloproteins. The essential requirement for a molecule to be considered a protein is that it feature one or more peptide (-HNC(O)-) bonds, as the amide hydrogen of the peptide bond (as well as those on the side chains of certain amino acid residues) has certain properties which permit it to be analyzed by hydrogen exchange.

"Agonist:" as used herein refers to a ligand that, upon binding to a receptor, triggers activation of a chemical signaling cascade that results in a definable change in the behavior or physical or biological state of a cell.

"Antagonist:" as used herein refers to a molecule that, by virtue of binding to a receptor or agonist, is able to block the cell-activating influence of the agonist, and which itself does not result in substantial activation of the cell.

"Labile Hydrogen:" is a general description of any hydrogen that may be exchanged with a solvent hydrogen under the conditions of the exchange experiment. One category of such hydrogens is the amide hydrogens found on the protein or peptide backbone. Exchange of these hydrogens, as described herein, is referred to as "amide hydrogen exchange". Another category comprises hydrogens bound to any carbon on the amino acids comprising the protein, including backbone and side chains that may temporarily be rendered labile under specific experimental conditions. Exchange of these hydrogens, as described herein, is referred to as "alkyl hydrogen exchange."

35

"Conditions of Slowed Hydrogen Exchange:" refers to conditions in which the rate of labile hydrogen exchange at residue atoms freely exposed to solvent is reduced substantially, *i.e.*, enough to perform the various manipulations necessary to handle the samples without undue or unacceptable loss of solvent-exposed labile hydrogens.

- 5 Conditions of slowed alkyl hydrogen exchange are easily, almost trivially achieved as alkyl exchange occurs only when particular experimental conditions temporarily induce it.

 Amide hydrogen exchange, however, is always occurring, to some extent, with a rate that is a function of factors including temperature, pH and solvent composition. The rate is
10 decreased three fold for each 10°C drop in temperature. Hence use of temperatures close to 0°C is preferred. In water, the minimum amide hydrogen exchange rate is at a pH of 2-3. As conditions diverge from the optimum pH, the amide hydrogen exchange rate increases, typically by 10-fold per pH unit increase or decrease away from the minimum. Use of high concentrations of a polar, organic cosolvent shifts the pH minimum to higher pH,
15 potentially as high as pH 6 and perhaps, with the right solvent, or even higher.

 By way of example, at pH 2.7 and 0°C., the typical half life of a tritium or deuterium label at an amide position freely exposed to solvent water is about 70 minutes. Preferably, the slowed conditions of the present inventions result in a half-life for a freely exposed amide hydrogen of at least 10 minutes, more preferably at least 60 minutes.

20

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 INTRODUCTION

 Many biological processes are mediated by noncovalent binding interactions between receptors and ligands. The identification of the structural features of the receptor
25 and ligand which immediately contribute to those binding interactions would be useful, *inter alia*, in understanding the mechanisms by which drugs bind to and activate receptors and in designing or identifying drugs which alter these processes.

 When both the receptor and ligand are proteins, the binding site of the pair may include as many as thirty amino acid residues on the receptor and an equal number on the
30 ligand. These residues constitute the "binding surfaces" of the respective members of the pair. However, the individual amino acid residues of the respective binding surfaces do not necessarily contribute equally to the free energy of binding of the complex. According to the studies of Wells and others discussed in the Background section, binding is at times and perhaps usually mediated through complementary "hot-spot" regions of a small number of
35 amino acid residues on the binding surfaces of each binding partner. Each of these "hot-

spot" residues makes a significant contribution to the free energy of binding, and in sum these relatively few residues can contribute as much as 40-50% of the free energy of binding of the complex. The hot-spot residues are surrounded by residues that contribute to the free energy of binding at lesser levels, and these residues are in turn surrounded by residues that
5 contribute at even lower levels.

As used herein, the term "hot-spot residue" is intended to include those residues of a receptor or ligand binding surface that individually contribute at least about 10% of the overall free energy of binding of the receptor-ligand complex. Preferably, each hot spot residue contributes about 20 to 30%, or even more, of the free energy of binding. In sum,
10 the hot-spot residues comprise a binding hot-spot for the receptor or ligand. Each binding surface may have a single hot-spot, or may have multiple contiguous or non-contiguous hot-spots.

Generally, the hot-spot residues *in toto* will constitute at least about 30%, preferably about 40%, and even more preferably 50% or more of the overall free-energy of binding of
15 the receptor-ligand complex.

Those of skill in the art will recognize that in many instances a ligand for a receptor may be a small polypeptide, or even a small organic molecule. In these situations, the receptor and/or ligand may not have a "binding surface." Indeed, in many of these instances, particularly those involving small molecule ligands, only relatively few receptor
20 amino acids might interact with the ligand. However, despite the lack of a true "binding surface," in these instances receptor and/or ligand hot-spot residues are as previously defined.

Hot-spot residues for the binding member of interest, which may be either the receptor, ligand or both, are identified according to the methods of the invention by
25 determining the hydrogen exchange rates of the individual labile hydrogens of the member in both the bound and unbound states, and correlating these measured exchange rates with contributions to the free energy of binding of the receptor-ligand complex. Because the exchange rates can be measured for a native receptor and/or native ligand under native, biologically relevant buffer conditions, the method of the invention permits, for the first
30 ~~time, the identification of hot-spot residues of a native~~ receptor or native ligand in its biologically relevant conformation.

Hydrogen exchange reflects the fact that many hydrogens in a macromolecule interchange or exchange with aqueous solvent hydrogens. Acidic hydrogens (O-H, N-H, S-H) readily exchange with aqueous solvent hydrogens under almost all physical and
35 chemical conditions. Most acidic hydrogens exchange too rapidly with solvent hydrogen to

be experimentally useful. By contrast, the more slowly exchanging peptide amide hydrogen exchanges slowly enough to be followed and measured. When contained within a polypeptide chain, every amino acid, whether gene-encoded or non-encoded, contains a peptide amide bond hydrogen, except for proline. Alkyl hydrogens (C-H) almost never
 5 exchange with aqueous solvent hydrogens except under specialised conditions and only then as long as the conditions persist. Alkyl hydrogen exchange can be induced by use of free radical generating agents including irradiation.

The measured exchange rate of a particular polypeptide labile hydrogen is directly related to the ease with which solvent water can gain close (atomic) proximity to the
 10 particular hydrogen, which is, in turn, dependant on the region of the polypeptide containing that labile hydrogen. Thus, measurement of the labile hydrogen exchange rates allow those hydrogens to be used as "sensors" or "reporters" of changes in the local environment surrounding any particular labile hydrogen. For example, in a system where a protein can be found in one of two allosteric states, the change in free energy of the local environment
 15 of a particular labile hydrogen when the protein switches between the two allosteric states is given by the formula:

$$\Delta\Delta G_i = -RT \ln(k_{i1}/k_{i2})$$

Equation (I)

20 where $\Delta\Delta G_i$ is the change in free energy in the environment of the i^{th} labile hydrogen when the protein switches from allosteric state 1 to allosteric state 2; R is the universal gas constant; T is the absolute temperature of the protein; ln is the natural logarithm; k_{i1} is the exchange rate constant for the i^{th} peptide labile hydrogen in allosteric state 1; and k_{i2} is the exchange rate constant for the i^{th} peptide labile hydrogen in allosteric state 2 (*see, e.g.*,
 25 Englander & Englander, 1994, Meth. Enzymol. 232:26-42; Bai et al., 1995, Meth. Enzymol. 259:344-356).

The atomic and molecular forces which determine polypeptide folding and conformational changes are identical to those that determine how two different proteins bind one another. Receptor-ligand binding surfaces are comparable to cross-sections
 30 through a single folded protein, with hydrophobic residues on the inside and hydrophilic residues on the outside. The receptor-ligand binding surface is held together by hot-spot residues. Around these hot-spots, the binding surface undergoes transient localized openings that expose parts of the surface to solvent. The differing protection factors for the several contact residues reflects their differential accessibility to solvent water in the
 35 flexible receptor-ligand surface.

Given the foregoing, a similar relationship applies to local changes in free energy upon formation of a receptor-ligand complex. In this instance, the ratio of the hydrogen exchange rate in the bound and unbound states for a particular labile hydrogen, or "protection factor" for that labile hydrogen, reflects the magnitude of the exchange slowing that labile hydrogen experienced upon formation of the receptor-ligand complex. The protection factor is related to the change in free energy between the bound and the unbound states at the location of that labile hydrogen in the receptor or ligand according to Equation (II):

$$\Delta\Delta G_i = -RT \ln(PF) \quad \text{Equation (II)}$$

where $\Delta\Delta G_i$ is the change in free energy in the environment of the i^{th} labile hydrogen upon formation of the receptor-ligand complex; "PF" or "protection factor" is the ratio of the exchange rates for the i^{th} peptide labile hydrogen in the bound and unbound states; and the other terms are as in Equation (I).

Equation (II) allows the free-energy contribution of each residue of the receptor to the overall free energy of binding to be determined. If every amino acid residue of a receptor's binding surface contributes equally to ligand binding (and vice versa), then the labile hydrogens at each residue of the binding surface would experience identical protection factors upon ligand binding. If, however, some residues of a binding surface are more important than others to binding, then the protection factors are not identical for each residue, and the labile hydrogens at the more contributory residues have protection factors that are higher than the average. Those residues with protection factors that are greater than the per residue average for the residues within the binding surface are the "hot-spot" residues of the invention.

Unlike an unfolding polypeptide, a receptor-ligand pair can completely dissociate. As a consequence, the fraction of time the receptor and ligand spend as a complex is a function of their concentrations in solution. In contrast, the folding behavior and the consequent labile hydrogen exchange rates of a single polypeptide is essentially independent of polypeptide concentration. The dissociative behavior of a receptor-ligand complex places an upper limit on the maximum protection factor that can be measured by the hydrogen exchange technique. Assuming equimolar amounts of receptor and ligand and a monovalent receptor-ligand complex, the maximum measurable protection factor is given by Equation (III):

$$MPF = ([R-L]/K_d)^{1/2} \quad \text{Equation (III)}$$

where MPF is the maximum protection factor; $[R-L]$ is the concentration of the receptor-ligand complex; and K_d is the dissociation constant of the receptor-ligand complex (Mayne et al., Biochemistry 31:10678-10685, 1992; Paterson et al., Science 249: 755-759, 1990). All protection factors in a binding surface that are above the MPF will be measured
5 as being equal to the MPF. Under the conditions usually employed in this technique, the maximum protection factor measurable in a binding interaction system with a K_d of 1×10^{-9} is 316; in a system with a K_d of 1×10^{-8} the MPF is 100, and in a system with a K_d of 1×10^{-7} the MPF is 31. If the affinity interaction is too low, hot-spot residues will not be distinguishable from less energetic regions of the binding surface. For a low affinity
10 interaction, the concentration of the receptor and the ligand should be as high as possible. Additionally, many low-affinity biological interactions are multivalent, thereby increasing the MPF for the system.

Thus, according to the methods of the invention, hot-spot residues for one or both members of a particular receptor-ligand pair of interest are identified by obtaining the
15 protection factors for each of the individual amino acid residues composing the respective receptor or ligand binding member and identifying those which contribute the requisite amount to the overall free energy of binding according to Equation (II).

As will be recognized from the above discussion, the methods of the invention rely on the recognition that labile hydrogen exchange rates in a polypeptide act as sensitive
20 reporters or sensors of local changes in the environment surrounding that particular labile hydrogen, and that those labile hydrogens which experience the greatest changes in exchange rates upon formation of a receptor-ligand complex reside in a region of the binding member that is thermodynamically more significant to the overall binding interaction than others. In this sense the labelling may be called "differential labelling"
25 because the extent of labelling of a particular site will be dependent upon its exposed (solvent accessible) surface area. Thus, while the expression "hot-spot residue" is used throughout the text to denote those amino acid residues in a receptor or ligand whose labile hydrogen has experienced a certain degree of change in solvent exchange rate upon formation of a receptor-ligand complex, the expression is not intended to imply that the
30 particular labile hydrogen itself, or the side chain of that residue, is necessarily involved in the binding interaction. The expression is merely used as a matter of convenience to identify those residues within the primary amino acid sequence of the receptor or ligand that reside in regions of the receptor or ligand that are in close proximity to thermodynamically significant binding interactions.

35

In order to determine the protection factors of the individual labile hydrogens of the member of interest, the hydrogen exchange rates of the individual labile hydrogens of the member in both the bound and unbound states must be determined.

In summary, measurement of receptor-ligand binding interaction exchange rates and protection factors for labile hydrogens, and, therefore, deduction of binding hotspots, requires:

1. An initial differential labelling, by hydrogen exchange, of the receptor and ligand at labile positions with heavy hydrogen isotopes in both the bound and unbound states.

2. Localisation and quantification of the heavy hydrogen label that has been attached to the receptor and ligand in both bound and unbound states.

The foregoing differential labelling and subsequent localisation and quantification steps can, with employment of appropriate experimental methodologies, be targeted to either amide hydrogens or alkyl hydrogens.

4.1 Differential Hydrogen Exchange Labelling of Receptor and/or Ligand

4.1.1 Amide Hydrogens in the Unbound State

For the unbound state, the individual amide hydrogen exchange rates can be measured during "off-exchange" or "on-exchange", or "on" then "off exchange" of heavy hydrogen using a variety of techniques. For example, exchange rates can be conveniently measured during on-exchange using conventional nuclear magnetic resonance (NMR) techniques, such as the techniques described in Cavanaugh et al., 1996, Protein NMR Spectroscopy Principles and Practice, Academic Press, New York, NY. In such an experiment, the member is contacted with a solution of deuterium or tritium ("heavy hydrogen"), preferably deuterium, under native conditions (typically conditions of physiological temperature, pH, ionic strength and buffer) and the rate of extinction of the amide hydrogen resonances followed by ^1H NMR. Alternatively, the member can be on-exchanged as a function of time and the quantities and locations of label determined according to the method described below.

For measurements during off-exchange, the member must first be labeled (on-exchanged) with deuterium or tritium by contacting the member with a solution of heavy hydrogen under the desired buffer conditions, preferably under conditions wherein the receptor adopts its native conformation, such as conditions of physiological temperature, pH, ionic strength, and presence of buffer salts. During this incubation period, heavy hydrogens from the solvent exchange with solvent-accessible peptide amide hydrogens,

thereby "on-exchanging" into the peptide residues of solvent-accessible portions of the member, including, among others, the peptide residues on the member's binding surface.

In some embodiments, the member is on-exchanged for a period of time sufficient to saturate the member with heavy hydrogen, i.e., for a period of time sufficient to label at least about 90%, and preferably about 95%, 96%, 97%, 98% or 99% or even more, of the amide hydrogens with heavy hydrogen. Typically, on-exchange periods on the order of a few to 24 hrs. when conducted at room or physiological temperatures and under physiological buffer and pH conditions (for example, 0.15mM NaCl, 10mM PO₄, pH 7.4; "PBS") are sufficient to saturate-label the member. Identification of an appropriate on-exchange period is within the capabilities of skilled workers. Preferably, the on-exchange is performed with small incubation volumes (0.1-10 μ l) and high concentrations of the member (10-100mg/ml). However, as will be discussed in more detail below, in other embodiments the member is on-exchanged as a function of time, typically by dispensing the member into a plurality of aliquots and on-exchanging the member in each aliquot for a different period of time or by removing aliquots from an on-exchanging sample at specified periods of time.

The on-exchanged member is then contacted with a solution of normal hydrogen under the same conditions of temperature, pH, ionic strength and buffer salts used for the on-exchange. Heavy hydrogens in solvent-accessible regions of the labeled member exchange with normal hydrogens in the solvent, thereby "off-exchanging" from the member residues at rates that depend on the solvent-accessibility of the residues.

The exchange rates of the individual amide hydrogens during the off-exchange period are measured using any of a variety of available techniques. For example, if the receptor or ligand is labeled with tritium, the rates of extinction of ³H NMR signals can be monitored and correlated to labile hydrogen exchange rates as previously described. Conversely, the build-up rates of ¹H NMR signals can be monitored. Similar measurements could be made with a deuterium-labeled receptor or ligand.

Alternatively, the individual amide hydrogen exchange rates can be determined by quantifying the amount of heavy hydrogen at each residue in the labeled member as a function of off-exchange time. In this embodiment, the receptor is first on-exchanged to saturation as described above and off-exchanged as a function of time, typically by dispensing the on-exchanged member into a plurality of aliquots and off-exchanging each aliquot for a different period of time or by removing aliquots from an off-exchanging member at specified time points. Typical time points are 1, 10, 100, 10³, 10⁴ sec., etc. Of course, the actual times used may vary, and may be experimentally determined for each system under study.

The off-exchange is quenched by changing the buffer conditions of the aliquots to conditions of slow amide hydrogen exchange, such as, for example pH 2.7 and 0°C. The locations and quantities of the heavy hydrogens are then determined for the member in each aliquot (time point) and this data analyzed using standard mathematical techniques to yield
5 the exchange rates of the individual labile hydrogens of the member.

4.1.2 Alkyl Hydrogen Exchange Labelling in the Unbound State

In a preferred embodiment, the hydrogen studied can be an alkyl hydrogen temporarily rendered labile by free radical mediated exchange with carbon bound hydrogens
10 within any portion of the side chain or backbone of any amino acid residue in the receptor or ligand. The mechanism by which exchange of such hydrogens is achieved is more complicated than the acid (H⁺) or base (OH⁻) catalysed amide exchange previously described. It is necessary to generate a potent hydrogen abstractor, in situ, to transiently “remove” hydrogens from the receptor carbon atoms. It is also necessary to ensure that a
15 source of heavy hydrogen (typically heavy water), and exchange catalyst is present that can replace the abstracted hydrogen (using solvent heavy hydrogen) and, in effect, repair the protein residue whose hydrogen had been removed.

In the preferred embodiment, the hydrogen atom abstractor is the hydroxyl radical and this is generated in situ by temporary application of radiolysis, for example by g-rays
20 from ¹³⁷Cs. (See, for example, M. B. Goshe, “Hydroxyl Radical Induced Hydrogen/Deuterium Exchange: Identification of Amino Acid Residues Involved in Peptide Protein Interactions”, Ph.D. Thesis, Case Western Reserve University, Ohio, 1999; M. B. Goshe and V. E. Anderson, “Hydroxyl Radical-Induced Hydrogen/Deuterium Exchange in Amino Acid Carbon-Hydrogen Bonds”, Radiation Research, 151, 50-58, (1999); M. B.
25 Goshe and V. E. Anderson, “Examining Hydroxyl Radical Induced Hydrogen/ Deuterium Exchange Into Peptides: Identification of Isotopically Modified Amino Acid Residues Using ESI-MS”, American Society for Mass Spectrometry, Annual Meeting, 1998.) As such, this mechanism may be finely controlled, in terms of the duration of irradiation (quantity or rate of hydroxyl radical generation), exchange facilitators (dithiotreitol, see
30 below) and factors that minimize adverse reactions (dinitrogen oxide). The amount of receptor or ligand and the amount of radiation must be titrated such that an amount of hydrogen exchanged product sufficient to give rise to a measurable signal in a mass spectrometer is produced. In practice, preliminary limited trial and error titration study is required for each sample of interest. Prior to irradiation, dissolved oxygen gas is removed
35 from the solution by bubbling dinitrogen oxide gas through it. Dinitrogen oxide performs a

dual function in that it also scavenges free electrons that arise from the irradiation. Upon radiolysis, or by some other means, hydroxyl radicals are generated in a concentration of, typically, 10mM. The hydroxyl radicals which are formed are highly reactive and abstract a hydrogen atom from the nearest carbon-hydrogen bond they can contact at the atomic level (a van der Waals radius). In practice such bonds are predominantly those found in amino acid residues when they are solvent accessible. Once a hydrogen atom has been removed, a carbon-centred radical remains. In order to ensure that the spare bonding position on this carbon becomes occupied by a heavy hydrogen, the receptor and/or ligand is present, at the time of irradiation, in a solution substantially enriched with the heavy hydrogen form of water. Furthermore, a heavy hydrogen donor, such as dithiothreitol (DTT), is simultaneously present in the solution where it, in effect, catalyses the exchange reaction. When DTT is used, the pH of the solution should be below 9.0. Due to the desire to obtain the greatest amount of heavy hydrogen exchanged product as possible, the heavy water should be present in high concentration, up to 80% mole fraction for deuterium and a specific activity of 50-100 Curies/millilitre for tritiated water. Once sufficient exchange has occurred as to allow subsequent measurement, irradiation is terminated, stopping further exchange and stopping undesirable side reactions.

In amide hydrogen exchange, labelling of receptor and/or ligand by on- and off-exchange either individually or sequentially can be experimentally useful. However, only on-exchange in heavy hydrogen containing buffers is useful in alkyl hydrogen exchange labelling. This is because it is anticipated that the experimental efficiency of alkyl exchange is substantially less than amide exchange, and the heavy hydrogen signals remaining after sequential on and off alkyl hydrogen exchange would be so low as to be useless. Consequently, in order to obtain exchange rates for alkyl hydrogens, one obtains only a timed series of on-exchanged samples of receptor.

Whereas with amide hydrogen exchange rate and protection factor determination, the time of on-exchange in solution is the critical independent variable, the variable important in determining alkyl exchange rates and protection factors is the irradiation time. Such irradiation times cannot be unduly long because there is the risk of damage to the receptor or ligand sample by undesirable radiation-induced reactions such as covalent cross-linking or formation of oxidation adducts. (Putting in simple terms, there is no concept of "saturation labelling" for alkyl hydrogen exchange because the time that would be required to achieve that is long enough to substantially degrade the receptor structure through side reactions.) Irradiation is best carried out for a period of time, sufficient to produce a reliable heavy hydrogen signal, as determined by preliminary titration experiments. In particular,

one may wish to focus attention, and therefore such titration studies, on a particular residue or side-chain alkyl hydrogen(s) of interest.

Typical irradiation times may comprise a sequence such as 1 minute, 10 minutes, 100 minutes. After the longest of these times, degradation of the protein structure may be sufficiently severe that one is only effectively going to obtain a lower limit on a protection factor. After irradiation, one carries out affinity chromatography on the products.

4.2 Localisation and Quantification of Exchanged Heavy Hydrogen Attached to Receptor or Ligand

The locations and quantities of the heavy hydrogens attached to the receptor or ligand by either amide or alkyl exchange can be determined using a variety of techniques, such as enzymatic and/or chemical decomposition followed by NMR analysis or radiation measurement or by mass spectrometry.

In the case of amide hydrogens, this is difficult because amide hydrogens are so labile that under many conditions they exchange with one another. Consequently, degradation of a receptor or ligand whose amide hydrogens have been labelled must be carried out in such a way that the labels do not move. By contrast, alkyl hydrogens are sufficiently unreactive (in the absence of free radical generating conditions) that, once a receptor and/or ligand has been labelled by attaching heavy hydrogens to carbon atoms, the labelled structure is stable to further change.

4.2.1 Localisation of Amide Hydrogens by Enzymatic Decomposition

In one embodiment, the quantities and locations of heavy hydrogens for each aliquot or time point are determined using the controlled sequential degradation methods described in U.S. Patent No. 5,658,739, application Serial No. 08/919,187 filed August 19, 1997 and/or WO 97/41436, each of which is incorporated herein by reference.

Briefly, in this embodiment, which has been developed for the special case of amide hydrogen exchange, the member is shifted to conditions of slowed hydrogen exchange, conditions which optionally simultaneously denature the member. The member is then cleaved into fragments of about 5-25 amino acids in length using any method amenable to conditions of slowed hydrogen exchange. Typical methods include, for example, chemicals such as CNBr or acid stabile proteolytic enzymes such as pepsin, Newlase, Aspergillus proteases, and protease type XIII, and mixtures thereof. Preferably, the member is fragmented with pepsin (10 mg/ml, covalently coupled to a solid support matrix such as Agarose or Poros® media (AL-20); at 0°C, pH 2.7, 0.1-20 min., preferably 10 min.)

The fragments are then separated from one another under conditions of slowed hydrogen exchange, for example by HPLC, and the labeled fragments identified, typically by radioactivity measurements, mass spectrometry or NMR (depending on the label) and isolated.

- 5 Each isolated labeled fragment is then sequentially degraded with, for example, carboxypeptidases, again under conditions of slow hydrogen exchange, to yield a series of subfragments wherein each subfragment in the series is shorter than the preceding subfragment by 1-5 amino acid residues, preferably by a single amino acid residue. As each carboxy-terminal amino acid of the off-exchanged member is sequentially cleaved by the
- 10 carboxypeptidase, the nitrogen which formed the slow-exchanging amide in the intact protein is converted to a rapidly-exchanging secondary amine, and any label at that nitrogen is lost from the protein within seconds, even at acidic pH. A difference in the molar quantity of label associated with any two sequential subfragments indicates that label is localized at the amide bond which differs between the two subfragments. Quantification of
- 15 the amount of label at each labeled amide, for example by radioactivity or mass spectroscopy measurements, as a function of varying on or off-exchange periods yields the exchange rate for each residue of the labeled fragments. The positions of the residues are then correlated within the primary amino acid sequence of the member to yield the exchange rates of the amide hydrogens of the member.
- 20 Preferred acid stable carboxypeptidases for the controlled sequential degradation include carboxypeptidases P, Y, W and C, with carboxypeptidase P being particularly preferred.

4.2.2 Localisation of Amide Hydrogens by Mass Spectrometry

- 25 In a preferred embodiment, there is a lessened reliance on sequential degradation by peptidases. Instead, fragmentation occurs within a mass spectrometer. Typically, when a peptide is fragmented within a mass spectrometer employing conventional conditions, amide hydrogens become scrambled, i.e., exchange positions with other amide hydrogens within the same peptide. But by performing specialised forms of mass spectroscopy (ion-
- 30 trap mass spectroscopy) and operating the mass spectrometer below a predetermined scrambling threshold fragmentation energy, such exchange can be avoided or minimised. (See, for example, Deng, Y., Pan, H., Smith, D. L. "Selective Isotope Labeling Demonstrates That Hydrogen Exchange at Individual Peptide Amide Linkages Can Be Determined by Collision-Induced Dissociation Mass Spectrometry", Journal of the American Chemical
- 35 Society; 121(9), 1966-1967, (1999); D. L. Smith, Y. Dheng and Z. Zhang, "Probing the

Non-covalent Structure of Proteins by Amide Hydrogen Exchange and Mass Spectrometry", Journal of Mass Spectrometry, 32, 135-146, (1997); D. L. Smith, provisional U.S. Patent application, "Methods of Quantifying Heavy Hydrogen Levels at Specific Peptide Linkages in Proteins and Peptides", Serial No. 60/113,772, filed 12/23/98; which are hereby
5 incorporated herein by reference). If this measurement is properly performed, the result is the generation of a series of subfragments similar in nature to the products produced by carboxypeptidases under conditions in which heavy hydrogen does not significantly interchange ("scramble") amongst the several amide hydrogen residues in the fragment. The weights of each subfragment are then determined within the mass spectrometer
10 allowing identification of subfragments with attached heavy hydrogen labels. Such conditions can be achieved with the use of an "ion trap" form of mass spectrometer whereby the energy of fragmentation can be carefully controlled. Consequently the resulting spectrum identifies precisely which residue contains the heavy hydrogen.

Conditions of slowed exchange must be employed prior to entry of the peptide to the
15 mass spectrometer apparatus. Prior to entry, an optional fragmentation of the receptor and/or ligand, with acid-reactive enzymes (pepsin, carboxypeptidase, etc.) can be performed. Alternatively an initial fragmentation can optionally be brought about by "collision induced dissociation" (CID) mass spectroscopy. Once the above procedure has been carried out for all fragmentation products obtained by the optional initial
20 fragmentation, the location of on-exchanged heavy hydrogens on the entire peptide sequence can be deduced. The operating principles described herein may also be achieved with other types of mass spectrometer, as would be obvious to one skilled in the art.

4.2.3 Localisation of Amide Hydrogens by Other Methods and

Summary

Alternatively, the controlled sequential degradation can be carried out using vapor phase pentafluoropropionic acid anhydride (PFPA) as described in Tsugita et al., 1992, Chemistry Letters pp. 235-238 and Tsugita et al., "Development of Novel C-Terminal Sequencing Methods," *In*: Methods in Protein Sequence Analysis, p. 55, Imahara &
30 Sakiyama, Eds., Plenum Press, New York, NY (1993), except that the protein is kept at -35°C for up to 3 hr. PFPA is then removed under vacuum, the subfragmented protein brought to 50 mM PO₄, pH 7.2 and analyzed by HPLC.

Those of skill in the art will recognize that in many instances the member will be of a sufficiently small size such that it need not be fragmented prior to the controlled
35 sequential degradation. Thus, in an alternative embodiment, the exchanged members are

optionally denatured under conditions of slowed hydrogen exchange and sequentially degraded as described above. Quantification of the amount of label at each amide as a function of exchange time yields the exchange rates for each residue of the unbound member.

5 Buffer conditions, protein concentrations and reagents, as well as automated procedures for carrying out each of the steps described above and additional optimal steps that can be used in conjunction with the invention are provided in U.S. Patent No. 5,658,739, application Serial No. 08/919,197 or WO 97/41436, supra, which are hereby incorporated herein by reference.

10 Alternatively, the quantities and locations of the heavy hydrogens for each aliquot or time point can be determined using NMR under conditions of slowed hydrogen exchange. An advantage of NMR is that the labeled receptor or ligand can be placed in an aprotic solvent which greatly slows hydrogen exchange, for example, solutions composed of greater than 90% dimethylsulfoxide (DMSO), acetone, chloroform, and other similar organic
15 solvents.

For these alternative methods, the whole receptor or ligand may be used, or optionally the receptor or ligand may first be denatured and fragmented into smaller pieces with, for example, chemicals such as CNBr or acid stable proteolytic enzymes, the labeled fragments identified and purified as previously described, and the locations and quantities of
20 heavy hydrogens within each fragment determined via radioactivity measurement (tritium exchange), mass spectroscopy or NMR (deuterium). Of course, it will be understood that the denaturing, fragmenting, purification and analysis must be carried out under conditions of slowed hydrogen exchange.

In general, the foregoing techniques of localisation may be employed in concert, if
25 so desired, once a receptor or ligand has been labelled. In one approach, optional preliminary fragmentation of the labelled receptor or ligand may take place if the molecule is large enough to require this, followed by an optional chromatographic separation of the fragmentation products. The fragmentation products may then themselves be subjected to further subfragmentation by collision induced mass spectroscopy, followed by examination
30 of the resulting spectrum.

4.2.4 Localisation of Alkyl Hydrogens

35

As previously mentioned, the process of identifying the positions of labelled alkyl hydrogens is rendered more convenient by the fact that such hydrogens are not inherently labile and do not readily exchange with one another.

The general principle that the labelled receptor or ligand is degraded into fragments, usually comprising a manageable number of amino acid residues, may be extended to complete degradation to individual amino acids. For example, this might be a sequential localisation by enzymatic or chemical digestion. At this level one aims to locate the position of the alkyl hydrogen labels to within one amino acid residue. The principle may be extended further to identify the precise location of the heavy hydrogen on the amino acid by further degradation into "subfragments", by mass spectrometry.

4.3 Exchange Rates for the Bound State

Next, the exchange rates of the individual labile hydrogens of the member in its bound state to a binding partner are determined. Like the exchange rates of the unbound member, the exchange rates of the labile amide hydrogens of the bound member can be determined during on-exchange or off-exchange, or on and off exchange in sequence, using the previously-described methods. Correspondingly, the exchange rates of the labile alkyl hydrogens in the bound state may be determined by on-exchange.

For both amide and alkyl hydrogens the foregoing methods of localisation and quantification are employed for the bound state in order to deduce the quantities of hydrogens labelled on residues of interest.

4.3.1 Amide Hydrogen Exchange in the Bound State

In general, the conditions employed to effect amide hydrogen exchange on bound states of receptor ligand complexes are just the same as those that would be employed for the unbound states.

For example, for on-exchange the amide hydrogen exchange rates can be conveniently measured via NMR. In one embodiment, the member is contacted with its binding partner under conditions conducive to binding, forming a receptor-ligand complex. ~~The receptor-ligand complex is then contacted~~ with a solution of heavy hydrogen, again under conditions conducive to binding, and the rate of extinction of the individual amide hydrogen resonances monitored by ^1H NMR. Since the amide hydrogens at the binding surface of the member are occluded from solvent by the presence of its binding partners, these amide hydrogens exchange reasonably slowly and have long-lived ^1H NMR resonances.

In another embodiment, the receptor-ligand complex is contacted with tritiated water under conditions conducive to binding and the build-up rates of the individual amide hydrogen resonances monitored by ^3H NMR.

For exchange rate measurements made during off-exchange, the member is first
5 labeled with heavy hydrogen as previously described. Again, the member is on-exchanged for a period of time sufficient to substantially saturate the member. The labeled member is then contacted with its binding partner under conditions conducive to receptor-ligand binding. Upon binding, the solvent accessibilities of those member heavy hydrogens that interact with the binding partner are significantly diminished by the bound partner. Hence,
10 the member heavy hydrogens that interact with the binding member are effectively "trapped" on the member.

The receptor-ligand complex is then dispensed into aliquots and each aliquot contacted with a solution of regular hydrogen under the same conditions used for formation of the receptor-ligand complex, and each aliquot off-exchanged for a different period of
15 time. During the off-exchange incubation, the heavy hydrogens in solvent-accessible regions of the receptor-ligand complex off-exchange from the receptor, typically at the same rate as they originally on-exchanged.

However, the heavy hydrogens incorporated in the binding surface of the member off-exchange with normal hydrogens from the solvent at a significantly slower rate, since
20 the bound binding partner decreases the solvent-accessibility of these residues. Again, off-exchange is quenched by placing the receptor-ligand complex under conditions of slowed hydrogen exchange, as previously described.

Typical time points for off-exchange are 1, 10, 100, 10^3 , 10^4 , 10^5 sec., etc., but again optimal time points for a particular binding pair can be readily determined empirically.

25 Following the off-exchange incubation, the receptor-ligand complex of each aliquot is dissociated under conditions of slowed hydrogen exchange, the desired binding member recovered and the quantities and locations of heavy hydrogen labels within the amino acid sequence of the member determined for the various off-exchange intervals as described above. Again, suitable conditions for carrying out the off-exchange and optimal steps that
30 can be used in conjunction with the methods are provided in U.S. Patent No. 5,658,739, application Serial No. 08/919,187 and/or WO 97/41436, *supra*, and are incorporated herein by reference.

Once the exchange rates for each labile hydrogen of the binding member of interest have been obtained in both the bound and unbound states, the hot-spot residue for that
35 particular member are identified according to Equation (II), *supra*. Those amino acids

whose labile hydrogen has a $\Delta\Delta G$ that constitutes at least 10-20% of the overall free energy of binding of the complex are the hot-spot residues for that particular binding member.

While the "protection factors" for each residue of the binding member of interest can be determined according to the methods described above, in preferred embodiments the
5 residues which constitute the binding surfaces of the member are first identified, preferably under native buffer conditions, and the protection factors of only those identified binding surface residues determined in order to identify the hot-spot residues. According to this preferred embodiment, the amino acid residues which constitute the binding surface of the member can be identified according to any available methods, including the methods
10 described in U.S. Patent No. 5,658,739, application Serial No. 08/919,187 and/or WO 97/41436; the methods of Clackson & Wells, 1995, *supra*, or from three-dimensional structural information for the receptor-ligand co-complex (where available).

In one convenient embodiment, the binding surface residues of the member are identified using the hydrogen exchange technique described in U.S. Patent No. 5,658,739.
15 During this process, a functionally-labeled member bearing heavy hydrogen labels only at residues which constitute the binding surface is generated, the functionally-labeled member is fragmented and the labeled fragments identified and isolated. Thus, by virtue of mapping the binding surface according to this method, conditions are obtained whereby fragments bearing labile hydrogens that reside within the binding surface of the member can be
20 generated and isolated. Fragmenting the on-exchanged or off-exchanged members obtained with the methods of the invention under these same fragmentation conditions followed by purification under the previously-identified isolation condition permits only fragments bearing binding surface labile hydrogens to be selectively isolated and analyzed for labeled content.

25 Alternatively, if three-dimensional or other structure-function information is available that identifies which residues constitute the binding surface for the member or co-complex, conditions for selectively isolating fragments which bear these residues can be determined, and then only these fragments from the off-exchanged member are isolated and analyzed for label content.

30 If the receptor-ligand complex can be dissociated without disrupting the native structure of the binding member of interest, a functionally labeled member can be obtained as described in U.S. Patent No. 5,658,739 and the exchange rates of the functionally labeled labile hydrogens in the unbound state obtained. The functionally-labeled complex is first dissociated under conditions of slowed hydrogen exchange, the desired binding member
35 isolated, and the labile hydrogen exchange rates of the functionally-labeled labile hydrogens

determined as previously described. For measurement of functionally labeled labile hydrogen exchange rates in the bound state, the functionally labeled co-complex is analyzed as previously described without first dissociating the complex.

- Other methods for measuring protection factors of only those labile hydrogens which reside within the binding surface of the member which can be used in conjunction with the methods of the invention will be apparent to those of skill in the art.

4.3.2 Preferred Method of Obtaining Amide Exchange Protection Factors

- In a preferred embodiment of the invention, the protection factors of only those labile hydrogens which reside in the binding surface of the member are conveniently obtained in a single experiment. In this embodiment, the member is on-exchanged as a function of time, typically by dispensing the member into a plurality of aliquots and on-exchanging the member in each aliquot for a different period of time, or by removing aliquots of on-exchanging member at defined time points. Typical on-exchange time points are 1, 10, 10^2 , 10^3 , 10^4 and 10^5 sec. Buffer conditions conducive to receptor-ligand binding should be used. The member in each aliquot is contacted with its binding partner and off-exchanged as a function of time, again typically by dispensing into aliquots. Each aliquot is off-exchanged for 0, 1, 10, 100 and 10^3 -times the on-exchange period. Preferably, the member is contacted with binding partner by passing it over a column of support-bound binding partner such as, e.g., binding partner-Sepharose. Following off-exchange, the reactions are quenched by changing the buffer conditions to conditions of slowed hydrogen exchange (typically pH 2.7 and 0°C) and the amounts and locations of label determined for the member of interest as previously described.

Performing the experiment in this manner yields the following matrix of data points:

	On Exchange Time (sec.)					
	1	10	100	10^3	10^4	10^5
Off Exchange Time (sec.)	0	0	0	0	0	0
	1	10	100	10^3	10^4	10^5
	10	100	10^3	10^4	10^5	10^6
	100	10^3	10^4	10^5	10^6	10^7
	10^3	10^4	10^5	10^6	10^7	10^8

- For each labile hydrogen in the member, plotting the quantity of label present for varying on-exchange times while holding the off-exchange time constant yields the

exchange rates for each labile hydrogen in the unbound state. Plotting the quantity of label present for varying off-exchange times while holding the on-exchange time constant yields the exchange rates for each labile hydrogen in the bound state. The protection factors for each labile hydrogen, and their respective contributions to the free energy of binding, are
5 determined as previously described.

Those of skill in the art will recognize that measuring protection factors according to this preferred method has several advantages. First, the matrix produces data redundancies which allow the quality of the measured exchange rates to be monitored. Data obtained for each off-exchange time (i.e., data obtained for diagonal matrix elements) yields exchange
10 rates for each labile hydrogen of the member in the unbound state. Data obtained for each on-exchange time (i.e. data obtained for column matrix elements) yields exchange rates for each labile hydrogen of the member in the bound state. Stated another way, each matrix diagonal provides exchange rates for each member labile hydrogen in the unbound state and each matrix column provides exchange rates for each member labile hydrogen in the bound
15 state. Thus, exchange rates across columns and diagonals can be compared as an internal check of the quality of the data.

Moreover, this matrix approach allows the exchange rates of labile hydrogens having different solvent accessibilities (and hence exchange rates) to be accurately determined. As previously discussed, labile hydrogens fall into several classes depending
20 on their solvent accessibilities: rapidly exchanging, slowly exchanging and those there between. The exchange rates of these different classes can be obtained from the data set which, by virtue of the time points bracketed, produces the highest quality data for each particular class of labile hydrogen.

In addition, the matrix approach permits the protection factors of only those member
25 residues of significance— those that reside in the binding surface— to be measured. This can be conveniently achieved by using a first off-exchange time point (t_1) greater than zero. Solvent accessible labile hydrogens, which do not reside in the binding surface, will exchange with solvent hydrogens and lose label during the period defined by $t_0 - t_1$. The actual time used to off-exchange these uninteresting labile hydrogens prior to collecting
30 time-dependent off-exchange data can be readily determined empirically. Typically an off-exchange pre-incubation on the order of from 1-5, 6, 7, 8, 9 or even as long as 10 seconds can be used to selectively unlabel labile hydrogens that do not reside in the binding surface of the member. By removing label from the rapidly exchanging solvent-accessible hydrogens, this method has the added advantage of allowing measurement of protection
35 factors for labile hydrogens of interest under optimal conditions of signal-to-noise. Only

those labile hydrogens of interest are labelled and are thus "seen" by the detection and quantification steps, thereby reducing background noise. Increases in signal to noise ratio over the two stage experiment of the order of 1,000:1 have been obtained.

While the methods of the invention have been described in terms of identifying hot-spot residues for one member of a receptor-ligand pair of interest, those of skill in the art will recognize that the methods can be used to identify hot-spot residues of both the ligand and receptor, either in individual experiments or simultaneously, so long as the receptor and ligand are each proteins of known amino acid sequence. Thus, the only requirement of the method is that the binding member for which hot-spot residues are desired be analyzable by hydrogen exchange, such as when the receptor or ligand is a protein of known or determinable amino acid sequence.

When hot-spot residues for both the receptor and ligand are desired, the exchange rates of the individual labile hydrogens for each binding member can be conveniently measured in a single experiment. For example, in instances when the NMR resonances of each binding member can be distinguished and assigned for the receptor-ligand complex, the exchange rates for the bound state can be conveniently measured in a single NMR on-exchange experiment.

Alternatively, if the NMR resonances of the complex cannot be resolved, the various off-exchange time points can be conveniently obtained in a single experiment by contacting a labeled receptor with a labeled ligand under conditions conducive to binding. After quenching off-exchange, the receptor-ligand complex can be dissociated and isolated under conditions of slowed hydrogen exchange and the exchange rates for the labile hydrogens of each binding member determined as previously described.

Preferably, the protection factors for each member of the binding pair are obtained in a single experiment using the data matrix approach described above.

4.3.3 Alkyl Hydrogen Exchange in the Bound State

Using similar principles as for amide hydrogen exchange, the method of labelling alkyl hydrogens for bound states is substantially the same as that for the unbound state. Only on-exchange in heavy hydrogen containing buffers need be employed. In order to obtain protection factors for alkyl hydrogens, one obtains only a timed series of on-exchanged samples of receptor bound to the ligand. From such measurements, exchange rates of receptor alkyl hydrogens in the bound state is deduced and protection factors calculated from the ratio for bound and unbound states. The same caveats with respect to

35

the times for which radiation can take place as were identified for the unbound receptor or ligand apply here too.

It is worth noting that in the case that the ligand is a small organic molecule bound to a protein receptor, the alkyl hydrogen exchange labelling can label alkyl hydrogens on that
5 ligand. Amide exchange labelling would, of course, only be able to provide direct information about the ligand contact regions if the ligand were a peptide or otherwise fortuitously contained an amide linkage.

10 **4.3.4 General Conditions Employed for Hydrogen Exchange in the Bound State**

For the various hydrogen exchange embodiments described herein that utilize radioactivity measurements and tritium labeling, the necessary level of tritiation (and hence the concentration of tritium in the buffer) is dependent on the total amount of protein available for analysis. For analysis of 1mg protein, at least 10 Ci/ml is desirable; for 0.1mg,
15 100 Ci/ml, and for .01mg, 1000 Ci/ml (pure tritiated H₂O is about 2500 Ci/ml.) For most applications, the tritiated water will be 50-500 Ci/ml. Without the use of these high specific activities, studies of proteins which are available in limited quantity would be much more difficult. Even higher specific activity (e.g., 500-1,500 Ci/ml) may be used in the invention, but radiation safety considerations necessitate performance of such on- and off-exchange
20 procedures in specialized facilities, such as are available in the tritium laboratory provided by the National Tritium Facility, Lawrence Berkeley Laboratories, University of California, Berkeley.

It should be noted that with customary levels of tritium, only a small percentage of the binding protein molecules will be tritiated at any given exposed position. All that is
25 required is that substantially each of the exposed labile hydrogen atoms be replaced in a number of the protein molecules that is detectable (by radiation counting).

For the various embodiments described herein utilizing deuterium labeling, on-exchange is preferably performed in 80-99% mole fraction deuterated water, more preferably 98% mole fraction deuterated water. The deuterated water may be optionally
30 supplemented with, for example, 2% mole fraction tritiated water (e.g., 50 Ci/ml). This modified procedure labels the binding member with both deuterium and tritium such that either radioactivity measurements, mass spectrometry and/or NMR techniques can be used to identify and isolate labeled fragments and/or to localize and quantify the positions of the label within the polypeptide sequences of the member. Methods for using radioactivity and
35 mass spectrometry measurements according to this modified embodiment are described in

U.S. Patent No. 5, 658,739, application Serial No. 08/919,187 and/or WO 97/41436 *supra*, and are incorporated herein by reference.

4.4 Combined Application of Techniques

5 In general, it should be noted that the foregoing steps may be combined with one another in appropriate orders where convenient. For example, it may be appropriate to identify residues of interest (receptor-ligand binding hotspots) via amide hydrogen exchange labelling, and then perform focused studies on just such regions with alkyl hydrogen exchange labelling. Iterative combinations of the steps may likewise be employed in order
10 to deduce both with convenience and specificity hot-spot residues.

4.5 Use of Hot Spot Information to Screen for Similar Binding Partners

In one extension of the foregoing discussion, it is apparent that the hydrogen exchange labelling and localisation techniques can be used to monitor changes in the
15 contribution of a residue to binding energy in a series of mutant proteins or polypeptides.

In another aspect, the invention provides methods of identifying compounds, particularly small organic compounds, that mimic the binding of a known ligand for a receptor of interest, *i.e.*, bind at or near the receptor hot-spot of a known ligand for the receptor of interest. As used herein, a ligand mimic is a compound which has a structure
20 different from the ligand but which interacts with at least one receptor hot-spot residue of the receptor-ligand complex (hereinafter referred to as "receptor_{R-L} hot-spot residues"). The greater the number of receptor_{R-L} hot-spot residues interacted with by the compound, the closer the compound is considered to mimic the ligand. Preferably, the compound will interact with at least a majority, or even more, receptor_{R-L} hot-spot residues. This approach
25 represents a way of probing "mutations" of small molecules in a conceptually similar way to the process behind comparing mutations of the receptor.

The overall procedure behind this approach is that protection factor measurement is used within an iterative scheme comprising three parts. In the first part, protection factors are measured, thus identifying hot-spot residues; in the second, families of related
30 molecules are obtained, for example by synthesis; and in the third part, those molecules are screened for activity against the receptor.

4.5.1 Iterative Discovery of Similar Binding Partners

According to the method of the invention, receptor_{R-L} hot-spot residues for a
35 receptor-ligand pair of interest are identified, typically under native conditions such as

conditions of physiological temperature, pH, ionic strength and buffer. As will be recognized by those of skill in the art, the method is not limited to any particular method of identifying receptor_{R-L} hot-spot residues. For example, the receptor_{R-L} hot-spot residues can be identified according to the previously-described methods of the invention, the site-
5 directed mutagenesis methods of Clackson & Wells, 1995 *supra*, or any other method described in the art. Due to its ability to identify receptor_{R-L} hot-spot residues for a receptor-ligand pair in its native conformation under native buffer conditions, the receptor_{R-L} hot-spot residues are preferably identified by the methods of the invention previously described herein. However, any other technique that quantifies the contribution to binding free energy
10 of individual receptor residues in a given receptor-ligand complex is sufficient to identify receptor_{R-L} hot-spot amino acids for the screening steps discussed below. In this context, an appropriate choice of ligand is one that is known to be a natural binding partner for the receptor in question, though any other ligand that has been found to bind well can be used, as can, in principle, any mutated form of the receptor.

15 For the purposes of determining hot-spot residues for use in screening experiments, one embodiment comprises carrying out single time-point measurements on amide hydrogen exchange. This approach enables such sites as can be to be saturation labelled thereby quickly providing lower limits on their protection factors.

After the receptor_{R-L} hot-spot residues have been identified, the receptor is screened
20 with candidate compounds to select those compounds that, by virtue of their binding to the receptor, slow hydrogen exchange at one or more of the receptor_{R-L} hot-spot residues. The candidate compounds can be any type of compound including biological compounds such as peptides, polypeptides, petidomimetics, nucleic acids, carbohydrates and non-biological compounds such as small organic molecules. The compounds can be synthesized
25 individually, or may be members of a library, such as a combinatorial library. As small organic molecules typically are preferred for pharmaceutical uses, these compounds are preferred for the screening steps described herein.

The candidate compounds can be screened directly for interaction with receptor_{R-L} hot-spot residues according to the methods of the invention, or can be first "pre-screened"
30 for a desired biological activity. For example, an initial screen can be performed by any conventional technique to assay for binding to the receptor, including non-selective binding assays and competitive inhibition of the binding of the native receptor-ligand complex. Other biological assays can also be used, and will be apparent to those of skill in the art, depending on the particular application.

35

Those candidate compounds that show some activity in the initial screen are then screened for interactions with the receptor_{R-L} hot-spot residues. The compounds are screened for interactions with a receptor_{R-L} hot-spot residue by receptor-candidate compound hot-spot screening, i.e., the hot-spot residues for the receptor-candidate
5 compound complex are identified (hereinafter referred to as "receptor_{R-C} hot-spot residues), typically using any of the previously described methods for identifying hot-spot residues. Those candidate compounds which produce a receptor_{R-C} hot-spot residue identical to a receptor_{R-L} hot-spot residue are considered to have interactions with a receptor_{R-L} hot-spot residue. Preferably, each candidate compound is screened as a ligand in the hydrogen
10 exchange techniques of the invention to identify the receptor_{R-C} hot-spot residues for the receptor-candidate compound complex. Most preferably, the protection factors of the receptor residues that interact with the candidate compound are measured by deuterium or tritium exchange. Alternatively, other techniques can be used to determine the receptor_{R-C} hot-spot residues for the receptor-candidate compound complex, including site-directed
15 mutagenesis and NMR, as previously described.

Those candidate compounds that interact with the receptor in a manner most similar to a ligand of the receptor are selected as "mimics" of the ligand. The receptor_{R-L} hot-spot residues are used to guide the selection of ligand "mimics" from the candidate compounds. Selection is based on the number of receptor_{R-L} hot-spot residues the candidate compound
20 interacts with, the magnitude of the protection factors of these particular receptor_{R-L} hot-spot residues for the receptor-candidate complex and the number of other receptor residues the candidate compound interacts with. Those compounds from the pool that interact with the greatest number of receptor_{R-L} hot-spot residues and the fewest number of other residues should be selected. Further distinction is provided by selecting those candidate compounds
25 whose receptor_{R-C} residue protection factors are similar in magnitude to the receptor_{R-L} residue protection factor of the same residue.

In practice, in an initial round of screening any candidate compound that interacts with at least one receptor_{R-L} hot-spot residues with any magnitude of protection factor is selected as a ligand mimic. This compound can then be modified, such as by acting as a
30 core or lead structure for the creation of a combinatorial library, and the modified compounds screened according to the methods of the invention. In subsequent rounds of screening, selection parameters should include binding more receptor_{R-L} hot-spot residues and providing protection factors closer in magnitude to those provided by the ligand. Eventually, compounds are isolated that closely mimic the ligand. Typically, these close

35

ligand mimics will precisely interact with a majority, and preferably all, of the receptor_{R-L} hot-spot residues.

Further selection criteria may include the ability of the candidate compounds to induce structural changes in the receptor, as identified by the ability of the candidate
5 compound to reproduce portions of the ligand-impressed thermodynamic fingerprint known to report such structural changes. In addition, the compounds can also be tested for their ability to function *in vitro* and/or *in vivo* as antagonists of the action of the ligand on the receptor.

The information learned from the above screening methods may be used in
10 conjunction with any other available data to guide the identification of compounds that precisely target the thermodynamic hot-spots of a given receptor-ligand complex. The methods can be used iteratively with successive rounds of hot-spot screening. In addition, other information, including three-dimensional structures of the receptor, ligand or the receptor-ligand complex, where available, can be used to guide the methods. However, it
15 will be recognized that success of the methods does not rely on the availability of such information—it can merely be used to enhance the methods when available.

4.5.2 Application to Combinatorial Chemistry

A candidate compound found to satisfy part of the thermodynamic fingerprint of the
20 receptor-ligand interaction as determined by hot-spot screening is in turn employed as the lead structure for the generation of a new combinatorial library of compounds from which additional compounds can be selected that act as ligand mimics. A combinatorial library, is really just a library of small molecule mutants. One approach is to have every member of such a library consist of this identified compound decorated with additional structural
25 features in a combinatorial manner, followed by selection of compounds that impress in the receptor not only the ligand's hot-spot fingerprint, but that also induce additional desired fingerprint interactions. Successive rounds of such combinatorial library generation and selection are performed until the desired level of affinity and/or specificity for the receptor is achieved. Progressive acquisition of the highest protection factor contact surface targets
30 in each round will result in isolation of compounds that, for their size, optimally interact with the largest number of energetically important receptor residues.

Promising compounds that target important receptor residues, but that also have an excessive number of irrelevant interactions with the receptor can be optimized by producing a combinatorial library that contains smaller versions of the compound with various
35 portions of the compound deleted and selecting compounds that retain reactivity with

thermodynamically important receptor residues but that no longer interact with unimportant receptor residues. Similarly, efforts to improve other properties of the compound, such as oral bioavailability, protein binding, and plasma half-life can proceed with the knowledge of which portions of the compound can be altered without a likely effect on the desired thermodynamic fingerprint. The concept of making purposeful modifications to an existing lead and assessing the results of such structural changes on functional activity is known in the art. It is an object of this invention to supplement the art by independently and simultaneously assessing the interactions of a compound with each residue of the receptor, or each residue of the receptor-ligand binding surface. Presently, the art is incapable of providing such detailed thermodynamic binding information absent solving the three-dimensional structure of the co-complex, which is a laborious, expensive undertaking. Even then, such detailed thermodynamic information can oftentimes only be inferred.

4.5.3 Application to Agonist Identification

The invention also encompasses an approach to identify compounds, particularly small organic compounds, with agonist activity from a pool of candidate compounds. A ligand's agonist activity towards a receptor is usually mediated by binding-induced oligomerization or conformational change in its cellular receptor.

The slowing of labile hydrogen exchange induced in a receptor by agonist ligand binding can arise by three distinct mechanisms:

(i) slowing of labile hydrogen exchange in residues that are in the receptor-ligand binding surface;

(ii) slowing of labile hydrogen exchange because particular residues participate in conformational changes as a consequence of the binding event (and therefore experience changes in their solvent accessibilities); and

(iii) slowing of labile hydrogen exchange in certain residues that participate in binding-induced receptor subunit oligomerization (and therefore experience decreased solvent accessibility because of these new subunit associations). If detailed information exists concerning the nature of the structure of the receptor-ligand complex, the simple inspection of such models in conjunction with the obtained protection factors of the residue allows discernment between amino acid residues in each of these three classes. The result is: (1) identification of residues that participate in receptor-ligand binding (and subsequent localization of receptor-ligand hot-spots); (2) identification of receptor residues that participate in conformational changes within the receptor; and (3) identification of oligomerization surface(s) of receptor sub-units.

If such structural information is unavailable, the receptor residues in the receptor-ligand binding surface can be identified and discerned from amino acid residues slowed from conformational change or oligomerization by labeling the ligand with heavy hydrogen according to the methods previously described, forming a receptor-ligand complex, and
5 identifying those receptor residues that become labeled by transfer from the labeled ligand. Methods for carrying out such label-transfer experiments are described in U.S. Patent No. 5,658,739, application Serial No. 08/919,187 and WO 97/41436, *supra*, and are incorporated herein by reference. Since this transfer occurs principally at the binding surface, the hot-spot residues on the receptor binding surface can be distinguished from
10 residues which exhibit changes in hydrogen exchange rates due to conformational changes and/or oligomerization. Those receptor residues whose hydrogen exchange rates change due to conformational changes and/or ligand-induced receptor oligomerization can be monitored in the methods of this invention to screen candidate compounds for potential agonist activity.

15 To identify compounds which exhibit agonist activity, a candidate compound is screened with the receptor as a ligand using the hydrogen exchange techniques described herein. Those candidate compounds that interact with at least one receptor residue belonging to class (ii) or (iii), *supra*, as determined by having a protection factor of at 1, are identified as having potential agonist activity, which is confirmed using standard biological
20 assays.

Alternatively, a pool of candidate compounds is first screened with the receptor to identify those candidate compounds that possess agonist activity. The identified compounds are then screened with the receptor using the hydrogen exchange techniques described herein to identify those compounds that interact with at least one receptor residue belonging
25 to class (ii) or (iii), *supra*. Iterative rounds of screening can be used to identify compounds which closely mimic the agonist, as determined by identifying compounds that interact with a majority of the class (ii) or class (iii) receptor residues, and optionally which induce protection factors substantially equal in magnitude at these residues to that induced by the agonist.

30 **As those of skill in the art will appreciate**, the methods of the invention are not limited to identifying compounds which mimic a receptor's natural ligand. In many instances, other ligands which bind a receptor are known. These ligands can be used in the methods of the invention to identify new ligands which mimic their binding and/or activity.

The foregoing written specification is sufficient to enable one of skill in the art to
35 practice the invention. Various modifications of the above-described modes for carrying

out the invention which are obvious to those of skill in the art are intended to be within the scope of the appended claims.

All cited references are hereby incorporated by reference in their entireties for all purposes.

5

10

15

20

25

30

35

What Is Claimed Is:

1. A method for identifying a hot-spot residue of a member of a receptor-ligand complex comprising the steps of:
 - 5 a. obtaining the protection factors of particular hydrogens on individual amino acid residues of which the member is composed; and
 - b. determining from the protection factors which individual amino acid residues of the member each contribute at least about 10% of the overall free energy of binding of the complex.
- 10 2. The method of claim 1 wherein said particular hydrogens are amide hydrogens on the peptide link between adjacent amino acid residues.
3. The method according to Claim 2 wherein said individual amino acid
15 residues comprise the binding surface of the member.
4. The method according to Claim 2 wherein those hot-spot residues identified in sum contribute about 20-30% of the overall free energy of binding of the complex.
- 20 5. The method according to Claim 2 wherein those hot-spot residues identified contribute about 40% of the overall free energy of binding of the complex.
6. The method according to Claim 2, in which the protection factors are obtained by a method comprising the steps of:
 - 25 a. measuring the hydrogen exchange rates of the amide hydrogens of the individual amino acid residues composing the member when the member is in an unbound state;
 - b. measuring the hydrogen exchange rates of the amide hydrogens of the individual amino acid residues composing the member when the member is in a bound
30 state; and
 - c. for each individual residue taking the ratio of the value obtained in step b with that of step a.

7. The method according to Claim 6 wherein the exchange rates of the amide hydrogens of the individual residues of the member in the unbound state are determined by a method which comprises the steps of:
- a. contacting the member with heavy hydrogen under conditions wherein
5 amide hydrogens in the member which are accessible to solvent exchange with, and become labeled by, heavy hydrogen;
 - b. contacting the labeled receptor with a solvent that is substantially free of heavy hydrogen under rapid exchange conditions for varying "off-exchange" periods during
10 exchange with normal hydrogen of the solvent;
 - c. for each "off-exchange" period, quantifying the amount of heavy hydrogen label associated with the amide hydrogen of each individual amino acid residue of the member; and
 - d. determining the hydrogen exchange rates for the amide hydrogens of
15 each individual member residue from the quantity of label associated with that residue as a function of the varying "off-exchange" period.

8. The method of Claim 7 wherein the amount of label associated with the amide hydrogens of each individual member residue for each off-exchange period is
20 quantified by a method comprising the steps of:
- a. fragmenting the labeled receptor into a plurality of fragments under conditions of slow hydrogen exchange;
 - b. determining which fragments are labeled with heavy hydrogen;
 - c. progressively degrading, under conditions of slow hydrogen exchange,
25 each of the heavy hydrogen-labeled fragments to obtain a series of subfragments, wherein each subfragment of the series is composed of about 1-5 fewer amino acid residues than the preceding subfragment in the series;
 - d. quantifying the amount of heavy hydrogen associated with each subfragment; and
 - e. correlating the amount of heavy hydrogen of the subfragments with the
30 amino acid sequences of the fragments from which the subfragments were generated, thereby localizing the positions of the fragment that had been labeled with heavy hydrogen to a resolution of about 1-5 amino acid residues.

35

9. The method of Claim 8 in which the heavy hydrogen is tritium and the presence or amount of heavy hydrogen on a fragment or subfragment is determined by radioactivity measurements.

5 10. The method of Claim 8 in which the heavy hydrogen is deuterium and the presence or amount of heavy hydrogen on a fragment or subfragment is determined by NMR spectroscopy.

11. The method of Claim 8 further including the step of separating the fragments
10 prior to determining which fragments are labeled with heavy hydrogen.

12. The method of Claim 11 in which the separation comprises two sequential separations steps which are carried out under different conditions.

15 13. The method of Claim 12 in which the first sequential separation step is carried out at a pH in the range of pH 2.1 to pH 3.0 and the second sequential separation step is carried out at a pH in the range of pH 2.1 to pH 3.0, where the pH of the first and second sequential separation steps are different.

20 14. The method of Claim 13 in which the first sequential separation step is carried out at a pH of 2.7 and the second sequential separation step is carried out at a pH of 2.1.

25 15. The method of Claim 8 in which the sequential degradation of the labeled fragments comprises contacting the labeled fragments with an acid resistant carboxypeptidase selected from the group consisting of carboxypeptidase P, carboxypeptidase Y, carboxypeptidase W and carboxypeptidase C.

30 16. The method of Claim 8 in which the sequential degradation of the labeled fragments comprises contacting the labeled fragments with pentafluoropropionic acid anhydride.

17. The method of Claim 8 in which any disulfide bridges in the off-exchanged member are disrupted under conditions of slow hydrogen exchange prior to fragmentation
35 or progressive subfragmentation.

18. The method of Claim 17 in which the disulfide bridges are disrupted by contacting the off-exchanged member or fragments with a water-soluble phosphine.

19. The method of Claim 8 where prior to fragmentation or subfragmentation,
5 the off-exchanged member is denatured under conditions of slow hydrogen exchange.

20. The method of Claim 19 in which the denaturation is carried out in a solvent such that the pH for minimization of hydrogen exchange is substantially higher than that of a purely aqueous solution.

10

21. The method of Claim 20 in which the solvent comprises about 5-20% water and the remainder is a nonaqueous polar solvent.

22. The method of Claim 21 in which the nonaqueous polar solvent is selected
15 from the group consisting of acetonitrile, dimethyl sulfoxide, a polyol and combinations thereof.

23. The method of Claim 22 in which the polyol is glycerol.

20 24. The method of Claim 21 in which the solvent further includes about 2-4 M guanidine thiocyanate.

25. The method of Claim 21, wherein the pH of the solvent is pH 4.8 - 5.2.

25 26. The method of Claim 6 in which the amide hydrogen exchange rates of the individual residues of the member in the bound state are determined by a method which comprises the steps of:

a. contacting the member with a heavy hydrogen-labeled solvent under conditions of rapid hydrogen exchange for an "on-exchange" period sufficient for solvent-
30 accessible amide hydrogens of the member to exchange with, and be substantially replaced by, heavy hydrogens of the solvent;

b. contacting the member with its binding partner under conditions wherein the heavy hydrogen labels are substantially retained and wherein the member binds its partner so as to form a receptor-ligand complex;

35

- c. contacting the receptor-ligand complex with a solvent that is substantially free of heavy hydrogen under rapid exchange conditions for varying "off-exchange" periods during which solvent-accessible amide hydrogens or heavy hydrogens of the receptor-ligand complex exchange with normal hydrogen of the solvent;
- 5 d. for each "off-exchange" period, quantifying the amount of heavy hydrogen label associated with the amide hydrogens of individual amino acid residues of the receptor; and
- e. determining the exchange rates of each individual member residue from the quantity of label associated with that residue as a function of the varying "off-exchange"
- 10 periods.

27. The method of Claim 26 wherein the amount of label associated with the amide hydrogens of the individual member residues for each off-exchange period is quantified by a method comprising the steps of:
- 15 a. dissociating the receptor-ligand complex and recovering the off-exchanged member of interest, both under conditions of slowed hydrogen exchange;
- b. fragmenting the off-exchanged member into a plurality of fragments under conditions of slow hydrogen exchange;
- c. determining which fragments are labeled with heavy hydrogen;
- 20 d. progressively degrading, under conditions of slow hydrogen exchange, each of the heavy hydrogen-labeled fragments to obtain a series of subfragments, wherein each subfragment of the series is composed of about 1-5 fewer amino acid residues than the preceding subfragment in the series;
- e. quantifying the amount of heavy hydrogen associated with each
- 25 subfragment; and
- f. correlating the amount of heavy hydrogen of the subfragments with the amino acid sequences of the fragments from which the subfragments were generated, thereby localizing the positions of the fragment that had been labeled with heavy hydrogen to a resolution about 1-5 amino acid residues.

30

28. The method of Claim 27 in which the heavy hydrogen is tritium and the presence or amount of heavy hydrogen on a fragment or subfragment is determined by radioactivity measurements.

35

29. The method of Claim 27 in which the heavy hydrogen is deuterium and the presence or amount of heavy hydrogen on a fragment or subfragment is determined by NMR spectroscopy or by measuring the mass of the fragment or subfragment.

5 30. The method of Claim 27 further including the step of separating the fragments prior to determining which fragments are labeled with heavy hydrogen.

31. The method of Claim 28 in which the separation comprises two sequential separations steps which are carried out under different conditions.

10

32. The method of Claim 31 in which the first sequential separation step is carried out at a pH in the range of pH 2.1 to pH 3.0 and the second sequential separation step is carried out at a pH in the range of pH 2.1 to pH 3.0, where the pH of the first and second sequential separation steps are different.

15

33. The method of Claim 32 in which the first sequential separation step is carried out at a pH of 2.7 and the second sequential separation step is carried out at a pH of 2.1.

20 34. The method of Claim 27 in which the sequential degradation of the labeled fragments comprises contacting the labeled fragments with an acid resistant carboxypeptidase selected from the group consisting of carboxypeptidase P, carboxypeptidase Y, carboxypeptidase W and carboxypeptidase C.

25 35. The method of Claim 27 in which the degradation of the labeled fragments comprises contacting the labeled fragments with pentafluoropropionic acid anhydride.

36. The method of Claim 27 in which any disulfide bridges in the off-exchanged member are disrupted under conditions of slow hydrogen exchange prior to fragmentation
30 or progressive subfragmentation.

37. The method of Claim 36 in which the disulfide bridges are disrupted by contacting the off-exchanged member with a water-soluble phosphine.

35

38. The method of Claim 27 where prior to fragmentation or subfragmentation, the off-exchanged member is denatured under conditions of slow hydrogen exchange.

39. The method of Claim 38 in which the denaturation is carried out in a solvent
5 such that the pH for minimization of hydrogen exchange is substantially higher than that of a purely aqueous solution.

40. The method of Claim 39 in which the solvent comprises about 5-20% water and the remainder is a nonaqueous polar solvent.
10

41. The method of Claim 40 in which the nonaqueous polar solvent is selected from the group consisting of acetonitrile, dimethyl sulfoxide, a polyol and combinations thereof.

42. The method of Claim 41 in which the polyol is glycerol.
15

43. The method of Claim 40 in which the solvent further includes about 2-4 M guanidine thiocyanate.

44. The method of Claim 40, wherein the pH of the solvent is pH 4.8 - 5.2.
20

45. The method of Claim 6 wherein the exchange rates for the amide hydrogens of individual amino acid residues of the member in the unbound state are determined by deuterium or tritium exchange NMR.
25

46. The method of Claim 6 wherein the exchange rates for the amide hydrogens of individual amino acid residues of the member in the bound state are determined by deuterium or tritium exchange NMR.

47. The method of claim 1 wherein said particular hydrogens are alkyl hydrogens on individual amino acid residues.
30

48. The method according to Claim 47, in which the protection factors are obtained by a method comprising the steps of:
35

- a. measuring the hydrogen exchange rates of the alkyl hydrogens of the individual amino acid residues composing the member when the member is in an unbound state;
 - b. measuring the hydrogen exchange rates of the alkyl hydrogens of the individual amino acid residues composing the member when the member is in a bound state; and
 - c. for each individual residue taking the ratio of the value obtained in step b with that of step a.
49. The method according to Claim 48 wherein the exchange rates of the alkyl hydrogens of the individual residues of the member in the unbound state are determined by a method which comprises the steps of:
- a. contacting the member with heavy hydrogen under conditions wherein alkyl hydrogens in the member which are accessible to solvent exchange with, and become labeled by, heavy hydrogen for varying "on-exchange" periods;
 - b. for each "on-exchange" period in step a., quantifying the amount of heavy hydrogen label associated with the alkyl hydrogen of each individual amino acid residue of the member; and
 - c. determining the hydrogen exchange rates for the alkyl hydrogens of each individual member residue from the quantity of label associated with that residue as a function of the varying "on-exchange" period.

50. The method of claim 8 in which the presence or amount of heavy hydrogen on a fragment or subfragment is determined by measuring the mass of the fragment or subfragment by mass spectrometry.

51. The method of Claim 8 in which the sequential degradation of the labeled fragments comprises application of collision induced mass spectrometry with a minimized scrambling threshold.

52. The method of Claim 8 in which the sequential degradation of the labeled fragments comprises a combination of enzymatic fragmentation and collision induced mass spectrometry.

53. A method of identifying a compound that binds a receptor at thermodynamically significant hot-spot residues bound by a known ligand for the receptor, comprising:

- 5 interest; and
- a. identifying the receptor hot-spot residues for the receptor-ligand pair of
 - b. selecting from a pool of candidate compounds those compounds that bind to at least one of the receptor hot-spot residues.

54. The method of Claim 53 wherein the candidate compounds are screened for
10 binding to the receptor and those compounds that bind are screened for binding to receptor hot-spot residues.

55. A method of Claim 53 in which the receptor hot-spot residues are identified by the method of Claim 1.
15

56. The method of claim 53 in which the receptor hot-spot residues are identified by the technique of site-directed mutagenesis.

57. A method of identifying a compound which acts as an agonist for a receptor
20 of interest, comprising:

- a. identifying for a receptor-agonist pair of interest those receptor residues which participate in (i) receptor-agonist binding; (ii) receptor conformational changes as a consequence of the receptor-agonist binding event (class ii residues); and (iii) receptor subunit oligomerization (class iii residues); and
- 25 b. selecting from a pool of candidate compounds those compounds that interact with at least one class ii or class iii receptor residue.

30

35

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/21177

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : Please See Extra Sheet.

US CL : Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/7.1, 23, 24; 436/34, 56, 57, 86, 87, 88, 89, 90, 173, 174, 175, 183, 501, 536; 549/26, 214, 388; 562/560, 561, 575

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST, EAST, DERWENT, STN

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,658,739 A (WOODS, JR.) 19 August 1997, see entire document.	1-57
Y	US 5,786,218 A (PIVONKA et al.) 28 July 1998, see entire document.	1-2, 6-10, 26-29, 45-46
Y	US 5,254,730 A (KILGORE) 19 October 1993, see entire document.	1-2, 6-10, 26-29, 45-49
A	US 5,101,059 A (CARPINO et al.) 31 March 1992, col. 2, line 53 to col. 20, line 12.	1-57
Y,P	WO 99/09204 A1 (WOODS, V.L.) 28 February 1999, see entire document.	1-57

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
B earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

18 NOVEMBER 1999

Date of mailing of the international search report

06 December 1999 (06.12.99)

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

GAILENE R. GABEL

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/21177

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (6):

G01N 33/53, 33/536, 23/00, 33/00; C12Q 1/37; C07C 229/00; C07D 311/82

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

435/7.1, 23, 24; 436/34, 56, 57, 86, 87, 88, 89, 90, 173, 174, 175, 183, 501, 536; 549/26, 214, 388; 562/560, 561, 575